

データベース

有次 正義

aritsugi@cs.kumamoto-u.ac.jp

目的(シラバスより)

- 情報化社会の基盤としてデータベースは不可欠な存在になっている。記憶装置の大容量化により、組織や個人が蓄積するデータ量は増大し続け、インターネットの普及は、そのようなデータの配布を容易にしている。本講義では、大量データを効率よく管理し必要な情報を簡単かつ高速に検索するデータベース管理システムの活用とその編成技法について学ぶ。

2

目標(シラバスより)

- 関係データベースの基礎理論について理解する
- 正規化について理解する
- 索引構造や障害時回復法について理解する

3

教科書・参考書

- 教科書:生協で購入してください
–「データベース入門[第2版]」
増永 良文著, サイエンス社
- 参考書
–「データベースシステム[改訂2版]」
北川 博之著, オーム社
–「リレーショナルデータベース入門[第3版]」
増永 良文著, サイエンス社



4

講義スタイル

- スライド(および板書?)
 - ノートを取る時間を考慮して進めます
 - スライドの進め方が早い時は, 手を挙げるなどして止めて下さい
 - ノート取りたかったのに取れなかったところは, 聞きに来てください
- 成績
 - 出席: 2/3以上なければ, 受験資格なし
 - 期末試験

5

スライド

- 以下から(以前使ったスライドに)アクセスできます

<http://www.dbms.cs.kumamoto-u.ac.jp/~aritsugi/Database/>

User: kumadai, passwd: 9mamo10

- 個人的な利用に限定してください!

6

定期試験

- 範囲は, 講義したすべての内容
- 1枚(A4サイズ)の片面手書きのメモの持ち込みを認めます
 - テストとメモの筆跡は合わせる
 - 学生番号・名前を必ず書く
 - 持ち込んだメモは, 必ず解答と一緒に提出してください. 採点することがあります

7

期待

- できるだけ講義時間内で理解する
- もしそれが無理でも, 次回の講義時間までに理解する
- できるだけコツコツ予習 & 復習をする

8

具体的には...

- 予習:教科書を読んでください
 - わかるところとわからないところを区別してください
 - わからないところを, 次の講義で理解できるようにしてください
 - わからないところを, 質問して理解できるようにしてください(相談が必要)
- 復習:講義の内容を見直してください
 - 教科書にない内容もいくつか話しますので, 講義の後に見直してください
 - わからない場合は, 質問して理解できるようにしてください(相談が必要)

9

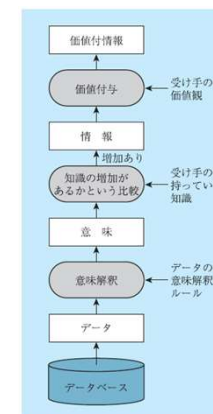
この講義の内容

- リレーショナルデータベース
 - データモデル
 - 問合せ
 - データベース設計
 - システム

10

第1章 データベースとは何か

データと情報の関係



データ (data) と情報 (information) は違うもの

データ: 記号
情報: 受け手の解釈

12

データベースとは

- コンピュータ内に構築された実世界の写絵
- 複数の応用目的での共有を意図して組織的かつ永続的に格納されたデータ群
- この講義では、データベースを作り、データを管理・利用するのに使われる、データベース管理システム(DBMS)の基礎について学ぶ

13

データと情報の関係

以下の二つを区別しよう

- データ
- 情報
- データ: 記号
- 情報: 受け手の解釈

意味のある(または、ない)なにか



知識とか経験とかに照らし合わせて

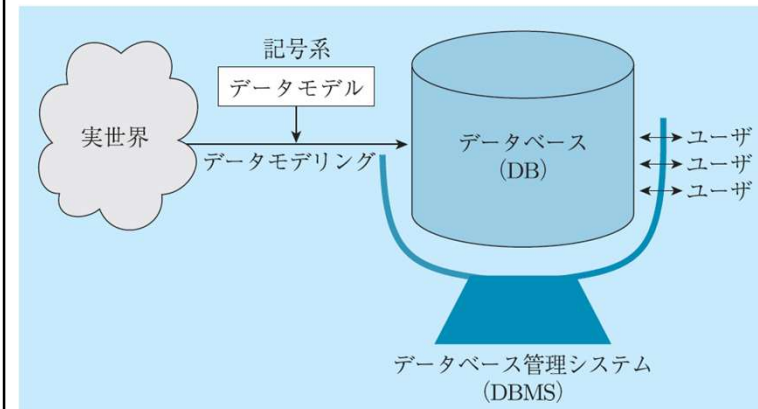


14

実現したいこと

- 複雑で大量のデータを高速に扱いたい
 - 所望のデータを簡単で高速に探し出したい: 検索
 - データの変更も高速に処理したい: 更新
- 複数のアプリケーション・利用者の間でデータを安全に共有したい
- データを使用しているアプリケーションが終了してもデータは保存しておき、再利用したい (データの永続性)

15



16

データモデル

- データベースは、コンピュータ内に構築された実世界の写絵
- 記号系, データモデルを使って実世界を写す
 - ネットワーク及びハイアラキカルデータモデル
 - リレーショナルデータモデル
 - オブジェクト指向データモデル
 - XMLデータモデル

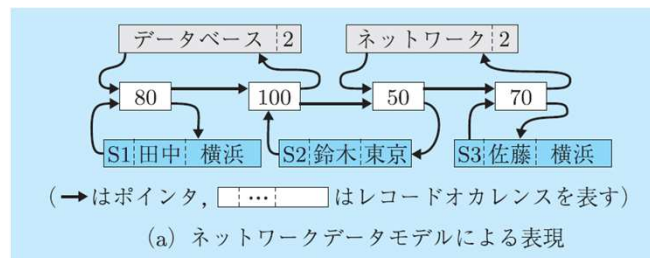
17

学生一履修一科目

- 学生が3人いる(田中, 鈴木, 佐藤)
- 科目は2つある(データベース, ネットワーク)
- 4つの履修関係があった

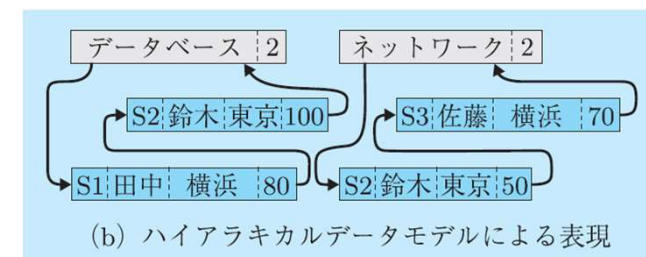
18

ネットワークデータモデル



19

ハイアラキカルデータモデル



20

リレーショナルデータモデル

科目		学生		
科目名	単位数	学籍番号	学生名	住所
データベース	2	S1	田中	横浜
ネットワーク	2	S2	鈴木	東京
		S3	佐藤	横浜

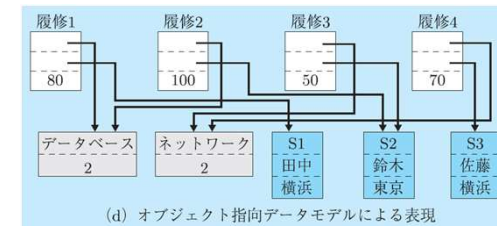
科目名	学籍番号	得点
データベース	S1	80
データベース	S2	100
ネットワーク	S2	50
ネットワーク	S3	70

(c) リレーショナルデータモデルによる表現

21

オブジェクト指向データモデル

- CADやマルチメディアデータなど、複合オブジェクトを管理するのに適したモデル
- オブジェクト指向のインターフェイスを備える



(d) オブジェクト指向データモデルによる表現

22

XMLデータモデル

```
<?xml version="1.0" encoding="Shift_JIS"?>
<履修 履修番号="1">
  <科目 科目名="データベース">
    <単位数>2</単位数>
  </科目>
  <学生 学籍番号="S1">
    <学生名>田中</学生名>
    <住所>横浜</住所>
  </学生>
  <得点>80</得点>
</履修>
...
<履修 履修番号="4">
  <科目 科目名="ネットワーク">
    <単位数>2</単位数>
  </科目>
  <学生 学籍番号="S3">
    <学生名>佐藤</学生名>
    <住所>横浜</住所>
  </学生>
  <得点>70</得点>
</履修>
```

(e) XML データモデル

23

関係モデル

- すべてのメジャーなデータベースシステムで用いられているモデル
- 数学の集合論に基づく
- とてもシンプル
- シンプルかつ十分な問合せの表現が可能
- システムの効率がよい

24

データベース管理システム(DBMS)とは

- (通常)OS上で動作するミドルウェア
- 主な機能
 - メタデータ管理機能
 - 質問処理機能
 - トランザクション管理機能

25

メタデータ管理機能

- メタデータ
 - データのデータ
 - 008040617,008040621,008040624, ...
 - なんのデータ？

26

メタデータ管理機能

- メタデータ
 - データのデータ
 - 008040617,008040621,008040624, ...
 - 最初の3ケタは教員番号で008は有次
 - 次の2ケタは令和何年
 - 次の2ケタは月
 - 次の2ケタは日
 - データの種類, サイズ(分布), インデックス情報, アクセス権, など

27

質問処理機能

- データベースに対する質問(query, 問合せ)を処理する機能
- 関係データベースシステムでは, SQLに代表される質問言語を使って, 質問を非手続き的(宣言的)に記述できる
- SQLで記述された質問の効率的な処理を実現する機能

28

トランザクション管理機能

- トランザクションとは、「データベースに対するアプリケーションレベルでの一つの原子的な作用」
 - 原子的(atomic):これ以上分解できない
 - 作用(action):データの読み(read)書き(write)
 - アプリケーションレベル:SQLの問合せ文と1対1に対応するわけではない
- ……今後の講義の中でまた出てきます
- 障害時回復
- 同時実行制御

29

DBMSが提供するものは..

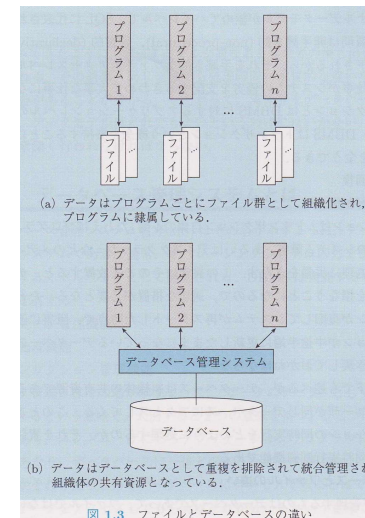
- 効率的で信頼性高く, 便利に, かつ複数人で使用可能な, 大量の永続データの保存と利用を可能にする機能

30

データベースとファイルの違い

- DBMSが管理するものは, データベース. ファイルではない.
 - データベース管理システム(DBMS)
 - Oracle Database, Db2, SQL Server, PostgreSQL, MySQL, HiRDB, ObjectStore
- データベースは共有資源. データを無矛盾に保つ.

31



32

ファイルだと困ること

- データとアプリケーションの相互依存
 - ファイルは(一つの)アプリケーションの中で使用. そのアプリケーションで使える形になっていればよい.
 - データベースは(複数のアプリケーションで)共有される
- 不十分な整合性維持機能
 - ある会社のサークルに所属する社員の社員番号が社員ファイルに登録のない社員番号となることはない, というようなことはファイルでは困難
- 不十分な機密保護
 - ファイル中の一部のデータのみを部外秘, といったきめ細かな制御はファイルでは困難

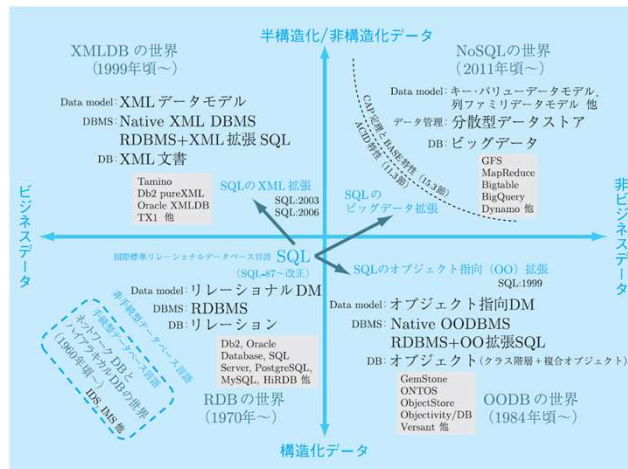
33

ファイルだと困ること

- 複数ユーザの同時アクセス
 - ファイルにもロックなどの機能はあるが, それを使って緻密にプログラムを作らなければならない
- 不十分な障害時データ保護
 - ファイルシステムは障害への対策はあまり考えられていない

34

データベースの潮流



35

まとめ

- データベースとは
 - 共有するデータを, 便利で安全に管理できる場所
- データモデルとは
 - 記号系
 - ネットワーク, 階層, 関係, オブジェクト指向, XML
- データベース管理システムとは
 - 効率的で信頼性高く, 便利に, かつ複数人で使用可能な, 大量の永続データの保存と利用を可能にする機能を提供するもの
- ファイルとの違い
 - DBは「共有する」というための機能がいろいろ満載
- Whether you know it or not, you're using a database every day hour.
- 次回は2章

36